PERSONAL EXPENSE TRACKER – DATA SCIENCE CAPSTONE PROJECT

ANALYZING AND PREDICTING PERSONAL SPENDING PATTERNS

Presented By : Aanchal Gor

EXECUTIVE SUMMARY

- Aim: Predict potential savings based on income, expenses, lifestyle.
- Dataset: Personal Expense Tracker with 19 columns.
- Methodology : Data Collection -> Data Cleaning -> EDA ->SQL Analysis -> Predictive Modeling -> Dashboard & Map
- Key Findings : Age , Income, City Tier affect savings ;Random Forest predicts well.

Presented By : Aanchal Gor

INTRODUCTION

•Problem: Individuals struggle to manage finances.

•Objective: Identify patterns & predict potential savings.

•Relevance: Helps budget planning.

•Placeholder: Optional infographic

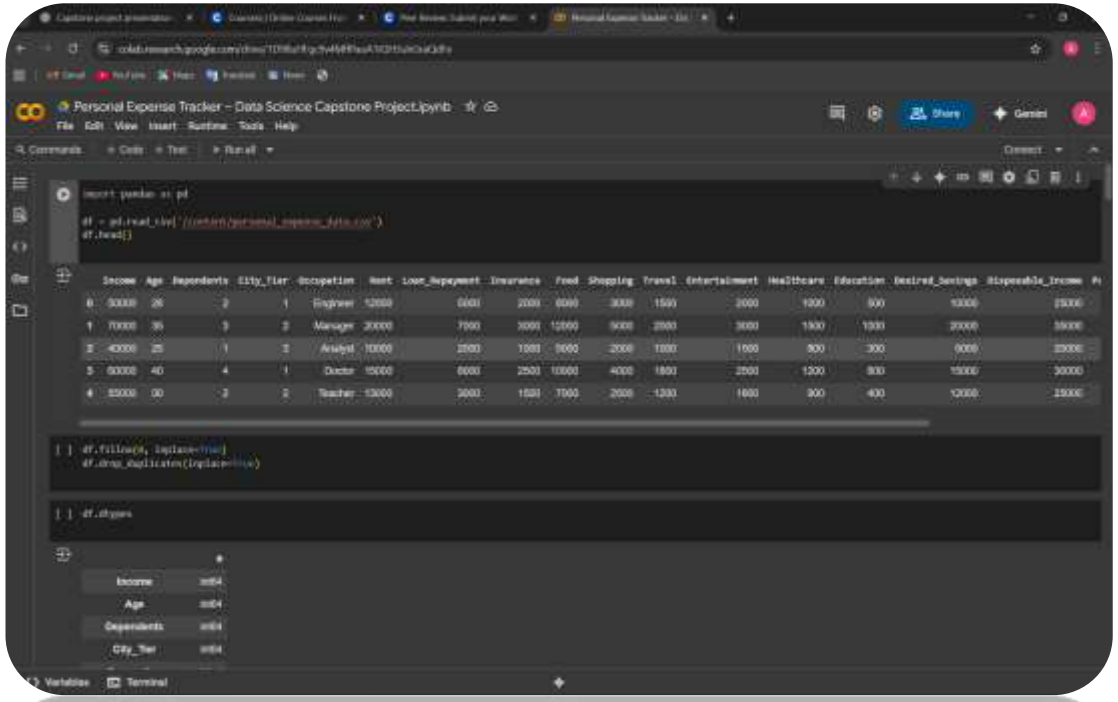
Presented By : Aanchal Gor

DATA COLLECTION & WRANGLING

•Data Source :- CSV FILE

•Cleaning :- Handle Missing Values, drop duplicates , convert categorical

•Placeholder: Table Screenshot of dataset head.



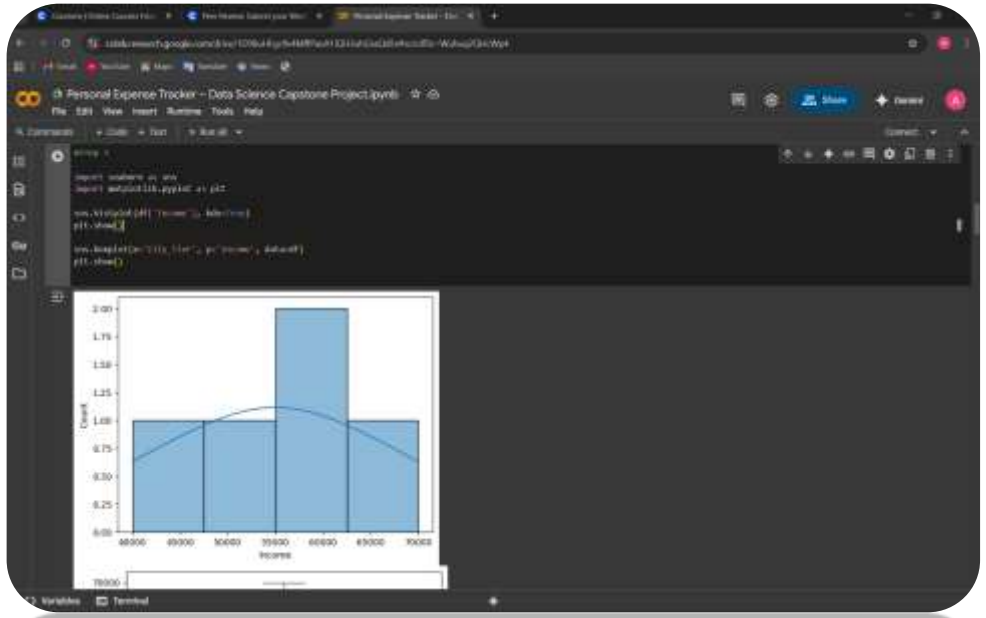
Presented By : Aanchal Gor

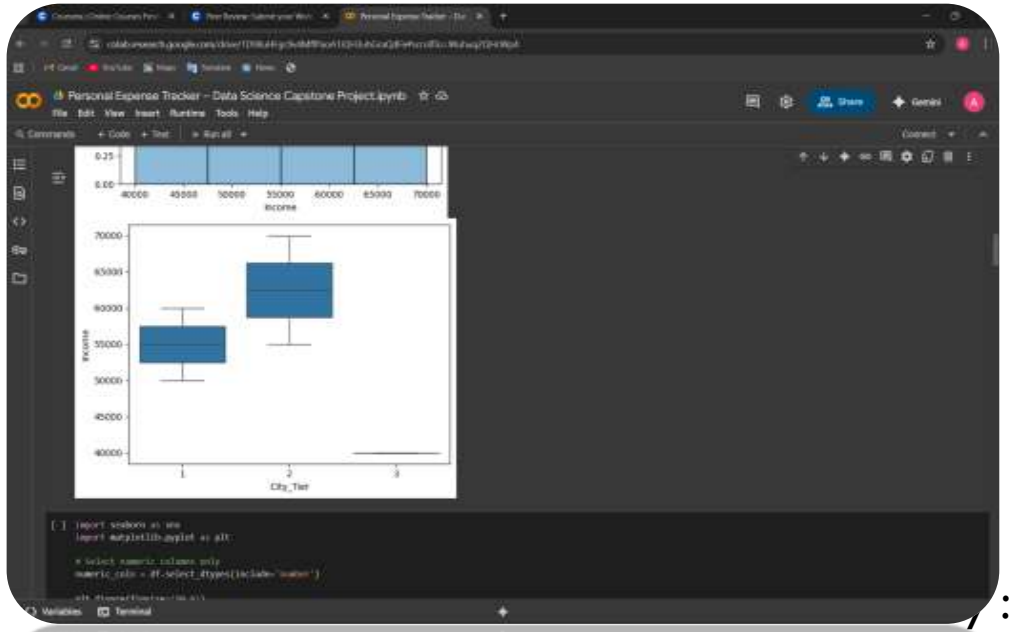
EDA & INTERACTIVE VISUAL ANALYTICS

•Techniques: Histograms, Boxplots, Correlation Heatmaps, Pairplots.

•Insights: Income higher in City Tier 1; Age correlates with savings; Rent & Loans affect Disposable Income.

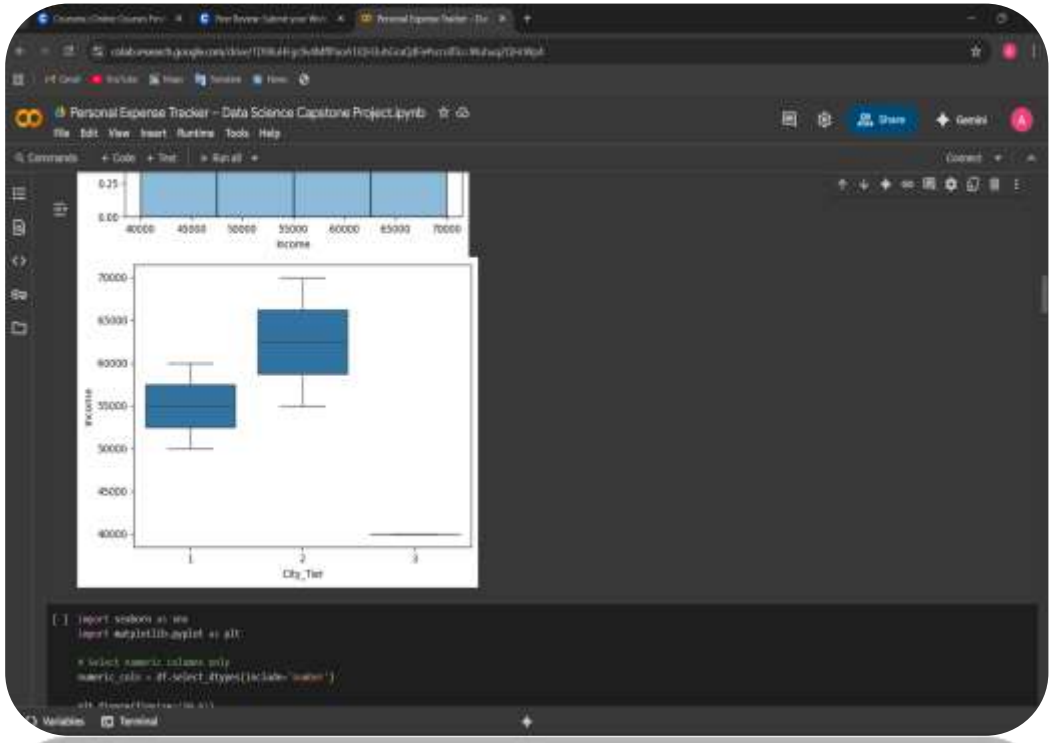
•Placeholder: Include heatmap, histogram, boxplot images.

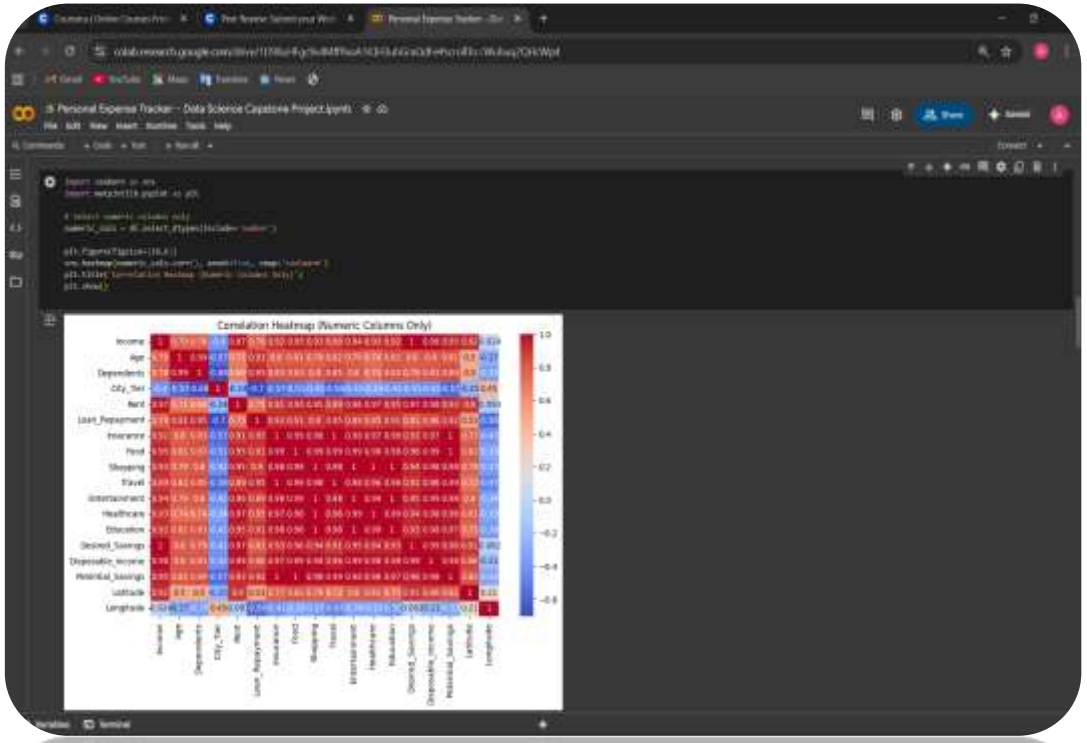




:Aanchal Gor

EDA & INTERACTIVE VISUAL ANALYTICS

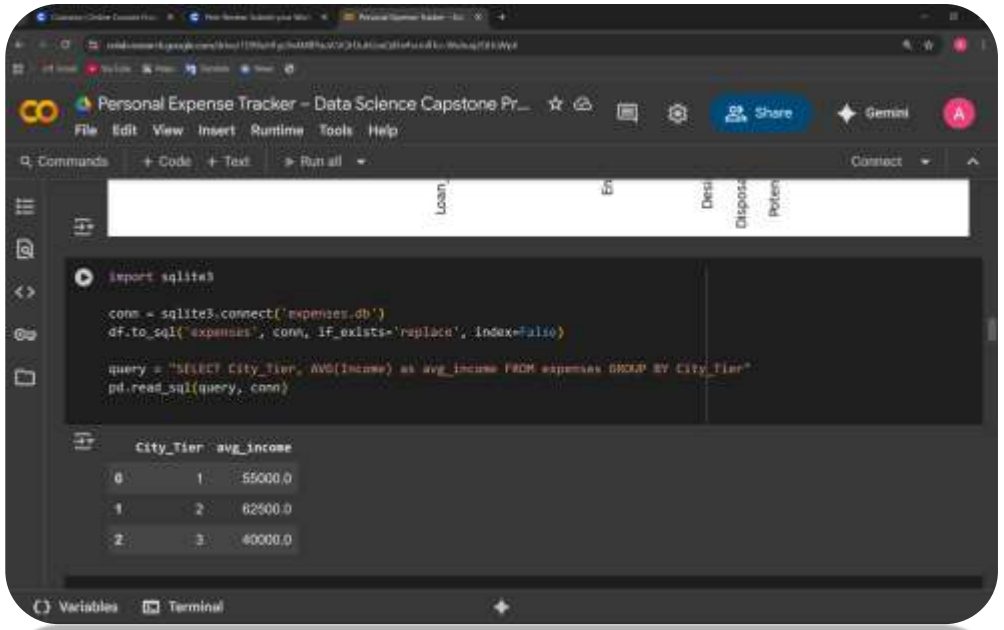


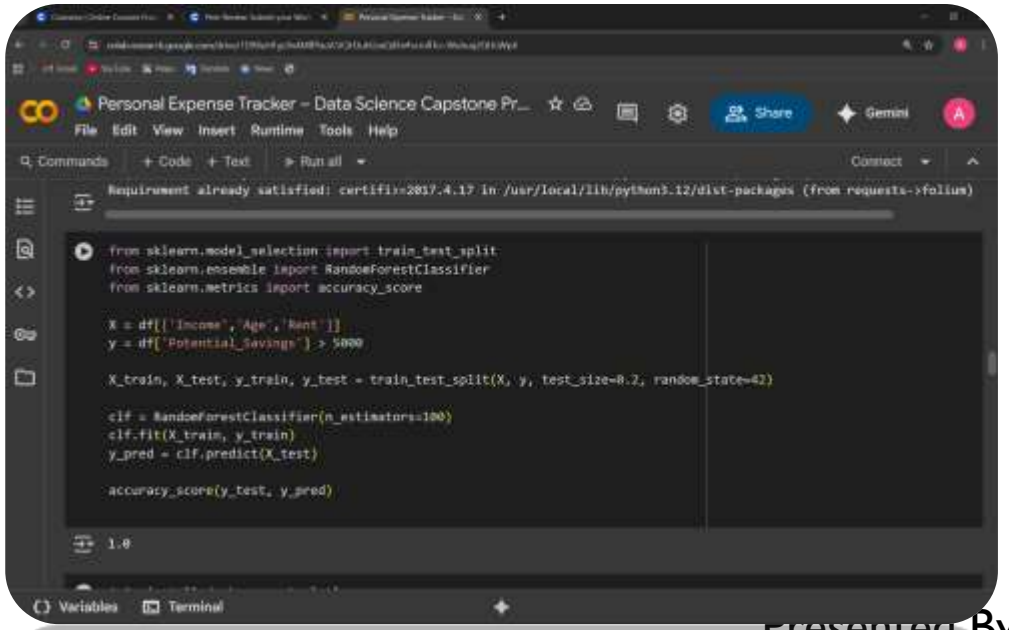


Presented By : Aanchal Gor

PREDICTIVE ANALYSIS METHODOLOGY

- Target: Potential_Savings > 5000 (binary).
- Features: Income, Age, Rent, Dependents
- Model: Random Forest
- Train/Test: 80% / 20%
- Performance metric: Accuracy & Confusion Matrix

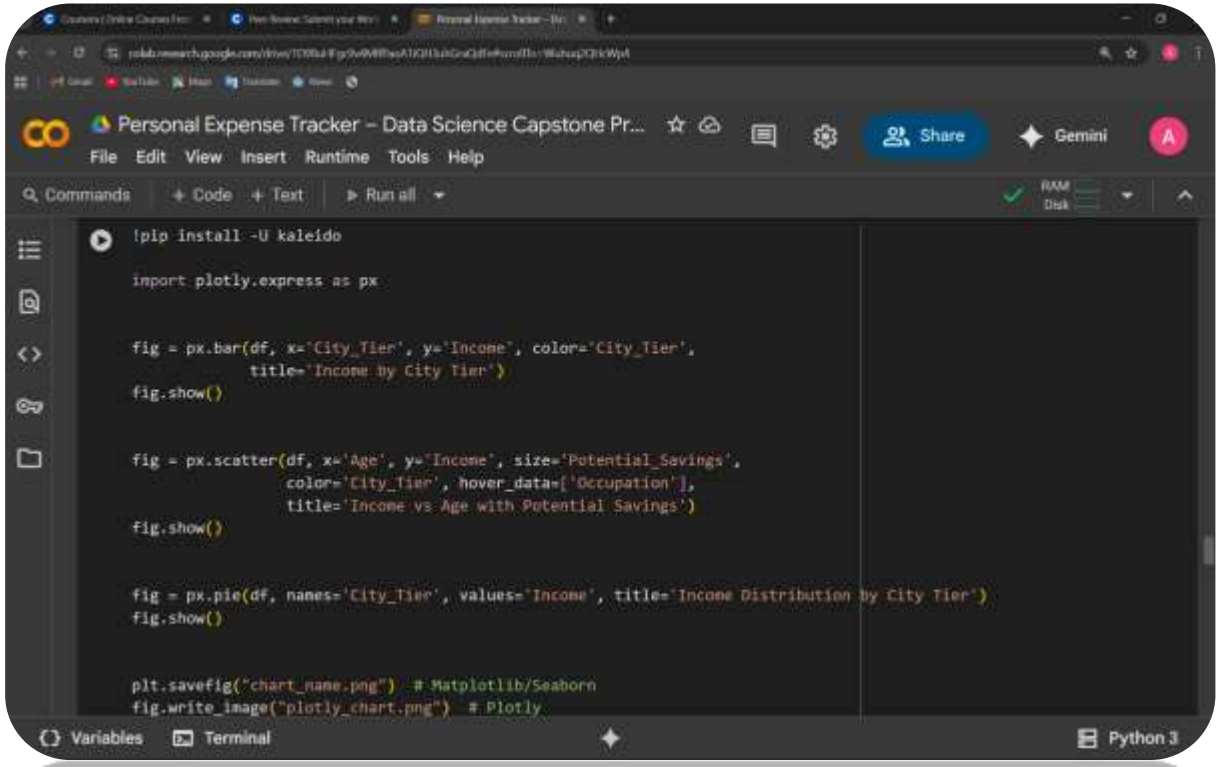




Presented By : Aanchal Gor

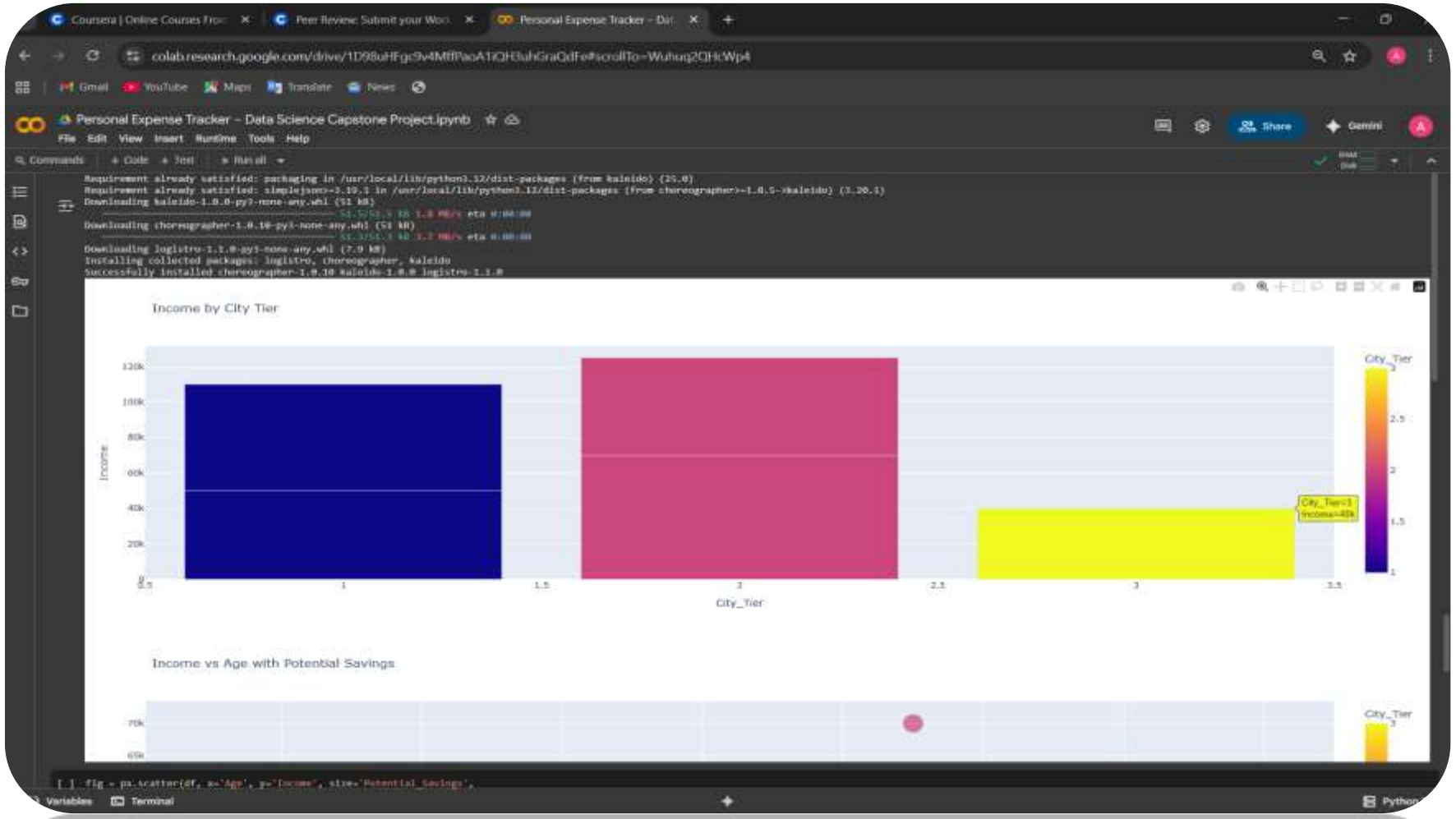
EDA WITH VISUALIZATION RESULTS

- Charts: Bar, Scatter, Pie, Pairplots
- Insights: Income distribution, spending patterns, correlations
- Placeholder: Chart screenshots



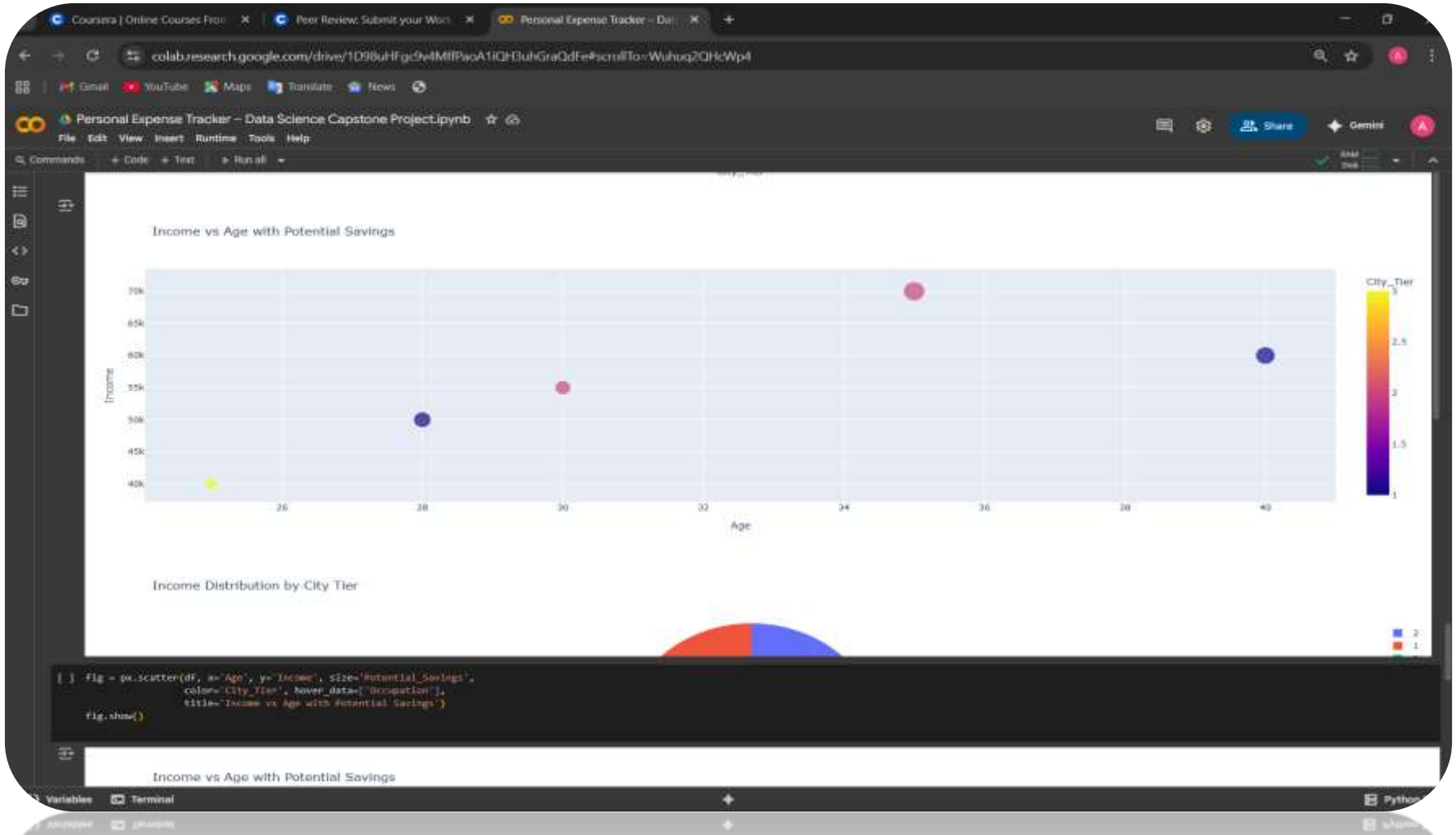
Presented By : Aanchal Gor

EDA WITH VISUALIZATION RESULTS



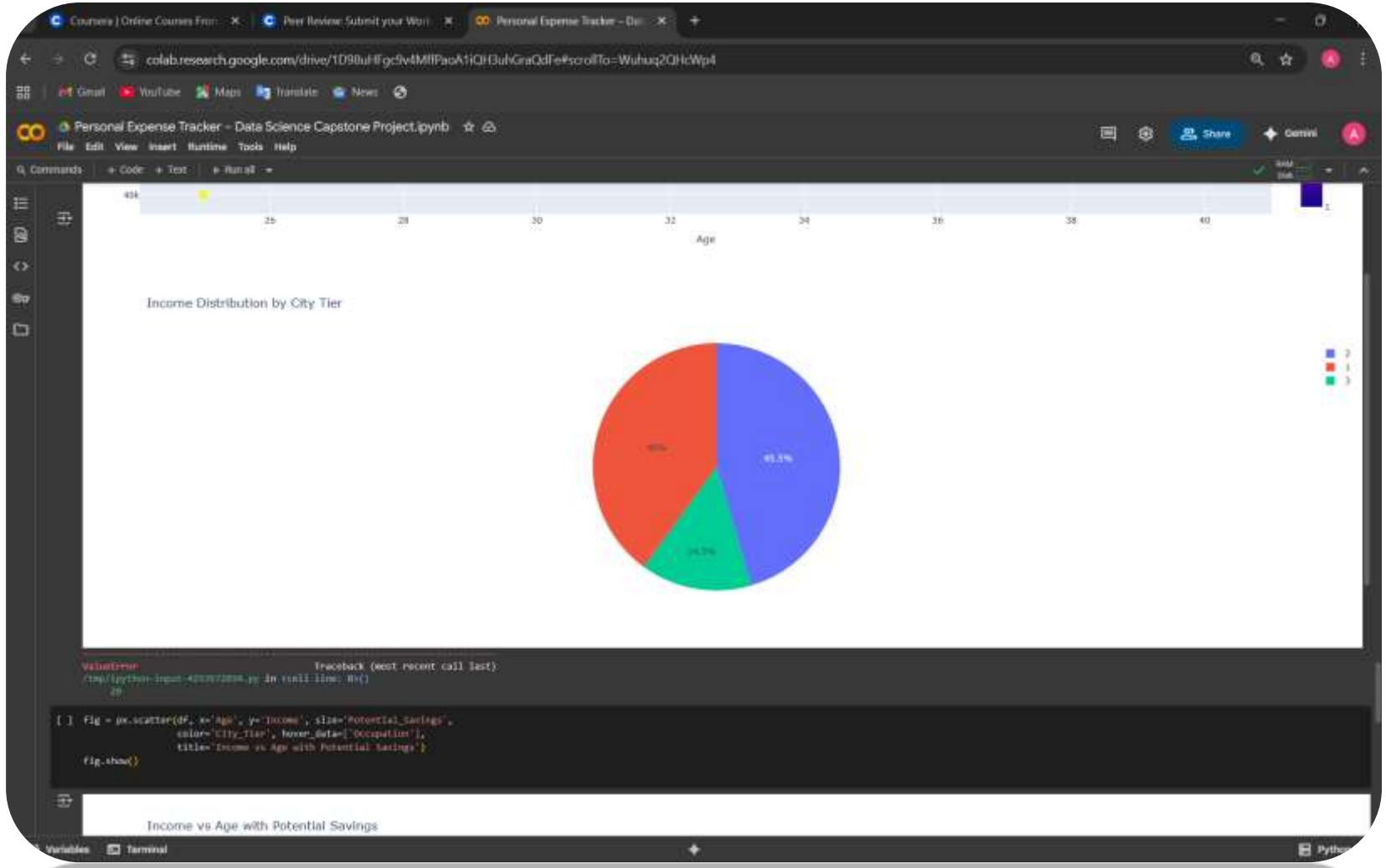
Presented By : Aanchal Gor

EDA WITH VISUALIZATION RESULTS



Presented By : Aanchal Gor

EDA WITH VISUALIZATION RESULTS



Presented By : Aanchal Gor

FUTURE WORK

Integration with SQL Databases

Automating data collection directly from SQL databases.

Creating optimized queries for faster retrieval and preprocessing.

Advanced Predictive Modeling

Applying machine learning models (Regression, Classification, Time Series).

Using historical SQL data to improve accuracy.

Scalability & Real-Time Predictions

Building real-time dashboards connected with SQL + ML pipelines.

Enabling live predictive insights for decision-making.

Data Visualization & BI Tools

Power BI/Tableau integration with SQL + predictive models.

Interactive visual reports for better understanding.

Continuous Model Improvement

Periodic retraining using updated SQL datasets.

Feedback loop for improving prediction performance

Presented By : Aanchal Gor

CONCLUSION

Key Findings :- Income, Age, Rent, City Tier influence savings; dashboards/maps help understanding

Future Work: Deploy web app, include more lifestyle variables

Presented By : Aanchal Gor

CREATIVITY & INNOVATIVE INSIGHTS

Extra Features: Interactive charts, Folium map, highlight patterns (e.g., students save less, doctors save more)

Future Work: Infographic of insights/trends

Presented By : Aanchal Gor

THANK YOU

Presented By :- Aanchal Gor